



McNair, A. G. K., Whistance, R. N., Forsythe, R. O., Macefield, R., Rees, J., Pullyblank, A. M., Avery, K. N. L., Brookes, S. T., Thomas, M. G., Sylvester, P. A., Russell, A., Oliver, A., Morton, D., Kennedy, R., Jayne, D. G., Huxtable, R., Hackett, R., Dutton, S. J., Coleman, M. G., ... Blazeby, J. M. (2016). Core Outcomes for Colorectal Cancer Surgery: A Consensus Study. *PLoS Medicine*, 13(8), [e1002071]. <https://doi.org/10.1371/journal.pmed.1002071>

Publisher's PDF, also known as Version of record

License (if available):  
CC BY

Link to published version (if available):  
[10.1371/journal.pmed.1002071](https://doi.org/10.1371/journal.pmed.1002071)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the final published version of the article (version of record). It first appeared online via Public Library of Science at <http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002071>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

GUIDELINES AND GUIDANCE

# Core Outcomes for Colorectal Cancer Surgery: A Consensus Study

**Angus G. K. McNair<sup>1,2,\*</sup>, Robert N. Whistance<sup>1,3</sup>, Rachael O. Forsythe<sup>1,3</sup>, Rhiannon Macefield<sup>1</sup>, Jonathan Rees<sup>1</sup>, Anne M. Pullyblank<sup>4</sup>, Kerry N. L. Avery<sup>1</sup>, Sara T. Brookes<sup>1</sup>, Michael G. Thomas<sup>3</sup>, Paul A. Sylvester<sup>3</sup>, Ann Russell<sup>5</sup>, Alfred Oliver<sup>5</sup>, Dion Morton<sup>6</sup>, Robin Kennedy<sup>7</sup>, David G. Jayne<sup>8</sup>, Richard Huxtable<sup>9</sup>, Roland Hackett<sup>10</sup>, Susan J. Dutton<sup>11</sup>, Mark G. Coleman<sup>12</sup>, Mia Card<sup>3</sup>, Julia Brown<sup>13</sup>, Jane M. Blazeby<sup>1,3</sup>**

**1** Centre for Surgical Research, School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom, **2** Severn School of Surgery, Bristol, United Kingdom, **3** Division of Surgery Head and Neck, University Hospitals Bristol NHS Foundation Trust, Bristol, United Kingdom, **4** Department of General Surgery, North Bristol NHS Trust, Bristol, United Kingdom, **5** Colorectal Consumer Liaison Group, National Cancer Research Institute, London, United Kingdom, **6** Academic Department of Surgery, University of Birmingham, Birmingham, United Kingdom, **7** Department of Surgery, St Mark's Hospital and Academic Institute, Harrow, United Kingdom, **8** Academic Surgical Unit, St James' University Hospital NHS Trust, Leeds, United Kingdom, **9** Centre for Ethics in Medicine, University of Bristol, Bristol, United Kingdom, **10** Colorectal Site Specific Group, Somerset, Wiltshire, Avon & Gloucestershire, South West Cancer Network, Bristol, United Kingdom, **11** Centre for Statistics in Medicine and Oxford Clinical Trials Research Unit, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom, **12** Department of Colorectal Surgery, Plymouth Hospitals NHS Trust, Plymouth, United Kingdom, **13** Clinical Trials Research Unit, University of Leeds, Leeds, United Kingdom



**OPEN ACCESS**

**Citation:** McNair AGK, Whistance RN, Forsythe RO, Macefield R, Rees J, Pullyblank AM, et al. (2016) Core Outcomes for Colorectal Cancer Surgery: A Consensus Study. *PLoS Med* 13(8): e1002071. doi:10.1371/journal.pmed.1002071

**Published:** August 9, 2016

**Copyright:** © 2016 McNair et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the MRC ConDuCT-II Hub (Collaboration and innovation for Difficult and Complex randomised controlled Trials In Invasive procedures - MR/K025643/1). RNW was supported by an NIHR doctoral research fellowship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** I have read the journal's policy and the authors of this manuscript have the following competing interests: AR received travel reimbursement from my home for attendance at Bristol University. AR is a patient member of the NCRI Consumer Forum and Primary Care Clinical Studies Group. AR is a patient member of the NCRI Colorectal Clinical Studies Group Adjuvant and Advanced Disease and Surgical Sub-groups. AR is a

\* [angus.mcnaair@bristol.ac.uk](mailto:angus.mcnaair@bristol.ac.uk)

## Abstract

### Background

Colorectal cancer (CRC) is a major cause of worldwide morbidity and mortality. Surgical treatment is common, and there is a great need to improve the delivery of such care. The gold standard for evaluating surgery is within well-designed randomized controlled trials (RCTs); however, the impact of RCTs is diminished by a lack of coordinated outcome measurement and reporting. A solution to these issues is to develop an agreed standard “core” set of outcomes to be measured in all trials to facilitate cross-study comparisons, meta-analysis, and minimize outcome reporting bias. This study defines a core outcome set for CRC surgery.

### Methods and Findings

The scope of this COS includes clinical effectiveness trials of surgical interventions for colorectal cancer. Excluded were nonsurgical oncological interventions. Potential outcomes of importance to patients and professionals were identified through systematic literature reviews and patient interviews. All outcomes were transcribed verbatim and categorized into domains by two independent researchers. This informed a questionnaire survey that asked stakeholders (patients and professionals) from United Kingdom CRC centers to rate the importance of each domain. Respondents were resurveyed following group feedback

member of the Independent Cancer patient Voice organisation.

**Abbreviations:** COS, core outcome set; COSMIN, Consensus-based Standards for the selection of health Measurement Instruments; COMET, Core Outcome Measures in Effectiveness Trials; CRC, colorectal cancer; RCT, randomized controlled trial

**Provenance:** Not commissioned; externally peer-reviewed

(Delphi methods). Outcomes rated as less important were discarded after each survey round according to predefined criteria, and remaining outcomes were considered at three consensus meetings; two involving international professionals and a separate one with patients. A modified nominal group technique was used to gain the final consensus. Data sources identified 1,216 outcomes of CRC surgery that informed a 91 domain questionnaire. First round questionnaires were returned from 63 out of 81 (78%) centers, including 90 professionals, and 97 out of 267 (35%) patients. Second round response rates were high for all stakeholders (>80%). Analysis of responses lead to 45 and 23 outcome domains being retained after the first and second surveys, respectively. Consensus meetings generated agreement on a 12 domain COS. This constituted five perioperative outcome domains (including anastomotic leak), four quality of life outcome domains (including fecal urgency and incontinence), and three oncological outcome domains (including long-term survival).

## Conclusion

This study used robust consensus methodology to develop a core outcome set for use in colorectal cancer surgical trials. It is now necessary to validate the use of this set in research practice.

## Background

Randomized controlled trials (RCTs) represent the gold standard in evaluating health care interventions. They aim to produce high quality evidence that can be used to inform clinical care; however, the clinical impact of RCTs is diminished by of a lack of coordination of outcome measurement and reporting. Indeed, multiple systematic reviews throughout many different branches of medicine have been consistent in demonstrating the large number and heterogeneity of outcome reporting in trials and other research studies [1–4]. This has the effect of making clinically relevant comparisons between trials and pooling of results in meta-analyses difficult. Furthermore, multiplicity of outcome measurement can lead to the selective reporting of significant findings in the form of outcome reporting bias [5].

A proposed solution to these issues is to develop and use “core outcome sets” (COSs). A COS is a minimum set of outcomes that key stakeholders agree to be measured in all trials in a particular field [6]. This approach allows a consistent set of outcomes to be measured and has the potential to improve the efficiency with which research can answer clinical questions. The benefits of COSs have now been embraced internationally by funding bodies [6], regulatory bodies [7,8], and journal editors [9], all of which recommend their use where available. As a result, the development of COSs is increasingly common. The COMET (Core Outcome Measures in Effectiveness Trials) initiative has recorded nearly 600 published or ongoing studies into COSs, and many have now been developed in diverse clinical areas including rheumatology [10], pediatrics [11], and obstetrics [12]. There is, however, no established COS for colorectal cancer (CRC) surgery.

This is now urgently needed, because CRC surgery is undergoing a period of intense innovation. CRC is a major cause of worldwide morbidity and mortality, representing the third most common cancer and fourth most common cause of cancer death [13]. Surgery is a fundamental method for both curative and palliative treatment of this disease, and there is therefore a great

need to improve the delivery of such care [14]. The last decade has seen several RCTs of laparoscopic techniques, all of which have measured different outcomes and thus suffer from the weaknesses described above [15,16]. The future will include evaluations of robotic surgery, transanal resection of the rectum, and organ-preserving rectal surgery, all of which have the potential to improve the care of many patients with CRC, provided they are evaluated in a robust and efficient manner. The aim of this study is therefore to define a COS for use in trials and other studies in CRC surgery, agreed upon by patients and CRC professionals.

## Methods

The scope of this COS includes clinical effectiveness trials (rather than trials of treatment efficacy) of all surgical interventions for cancer of the colon and rectum. Excluded were oncological interventions. The COS defines which outcomes are recommended, but does not specify how they should be measured. The COS could also be used in audit and nonrandomized studies of CRC.

The development of the COS was conducted in three phases according to COMET guidelines [6]. In Phase 1, a long list of outcomes that could be measured in CRC trials was identified, and outcomes were categorized into domains. In Phase 2, domains were operationalized into a questionnaire that was used to survey stakeholders' views on the importance of each domain using Delphi methods. In Phase 3, consensus meetings with patients and surgeons were used to finalize the core set. Appropriate ethics regulatory approval was granted (National Research Ethics Service number 10/H0102/82).

### Phase 1: Domain Generation

Outcomes of CRC surgery were identified from three sources; i) systematic review of clinical and patient-reported outcome literature [17,18]; ii) interviews with patients; and iii) analysis of written patient information leaflets used for colorectal surgery in hospitals in the United Kingdom. Duplicates were removed, and a long list of outcomes was created. Similar outcomes were categorized into domains by two members of the study team. Patient-reported outcomes were grouped into domains (e.g., ability to walk and activity levels were grouped within the physical function domain) and verified by two researchers and a patient representative [19]. Items from patient information leaflets were independently categorized by two surgeons. Discrepancies were resolved through discussion with the study lead. Overlapping domains between data sources were condensed, producing a final list of domains.

The final domains were operationalized into questionnaire items using lay language with the medical terminology included in parentheses. The questionnaire was piloted by patients for face validity, understanding, and acceptability and modified as a result of this feedback.

### Phase 2: Delphi Consensus Methods

The questionnaire developed in Phase 1 was sent to stakeholders including CRC surgeons, clinical nurse specialists, and patients who had undergone surgery for CRC (Round 1). Patients were considered to be essential stakeholders, as they are the recipients of treatment, and surgeons and clinical nurses have an in-depth understanding of the potential impact of surgery. Oncologists were excluded as chemo/radiotherapy was outside the scope of this COS. Surgeons and nurses were identified from United Kingdom (UK) National Health Service hospital trusts that routinely performed surgical resection of CRC and participated in the UK National Bowel Cancer Audit. Nonprobabilistic purposive sampling was conducted to ensure center variation based upon geographical region (Northern England, the Midlands, Southwest and Southeast England, and Wales) and caseload volume per annum as determined by

number of major resections in 2012. Patients were recruited from University Hospitals Bristol NHS Foundation Trust, North Bristol NHS Trust, and Plymouth Hospitals NHS trust in the UK. Participants were approached by post and were sent the questionnaire with a stamp-addressed envelope for return. One reminder was sent if there was no response after four weeks. Nonprobabilistic purposive sampling was conducted to ensure representation based on age, sex, and cancer site (rectum, left colon, right colon). Demographic data was collected including area of deprivation, marital status, employment status, and educational level. Deprivation was defined by the UK Office of National Statistics Index of Multiple Deprivation at lower layer Super Output Area level for the individual. This is a combined measure of income, employment, health and disability, education, barriers to public services, crime, and living environment. Educational level was defined as up to basic education (to the age of 16 or completion of the UK General Certificate of Secondary Education or equivalent), further education (subsequent qualifications to the age of 18 but not degree level), undergraduate, and postgraduate education.

Questionnaires asked participants to rate the importance of domains on a nine-point Likert scale, where 1 was a “not essential” and 9 an “absolutely essential” outcome. Returned first round questionnaires were analyzed, and any outcomes considered least essential were discarded. In Round 2, participants were provided with feedback from Round 1 in the form of their previous score for each domain and a mean score from their stakeholder group. Participants were then asked to rescore each domain on the nine-point Likert scale, and the results were used to determine which domains should be retained and presented in the consensus meetings. Participants that did not respond to the first questionnaire were ineligible for Round 2 because of the necessity to receive their own feedback. Responses from Round 1 were accepted until the Round 2 questionnaire was distributed. Round 2 responses were accepted until the respective stakeholder consensus meeting.

### Phase 3: Face-to-Face Consensus Meetings

Three consensus meetings were held; two with health professionals and a third with patients and caregivers. The first professional consensus meeting was held at the Tripartite Colorectal Meeting (the combined meeting of the Association of Coloproctology of Great Britain and Ireland, the American Society of Colon and Rectal Surgeons, the Royal Society of Medicine, Royal Australasian College of Surgeons, Colorectal Surgical Society of Australia and New Zealand, and the European Society of Coloproctology) in Birmingham in 2014. Ongoing discussion prevented the completion of the consensus meeting within the allotted time, and a second was hosted by the European Society of Coloproctology meeting in Barcelona in 2014. Meetings were open to all members of international societies and, in addition, all participants of the Delphi process were invited to attend. Participants were asked to declare their country of residence. The patient and caregiver meeting was held in Bristol in 2013. Attendees at this meeting were all from the UK and had completed the questionnaire surveys and responded to an invitation to attend a consensus meeting.

The retained outcomes from the second survey were presented at the meetings, and participants were asked to anonymously rate their importance. Anonymized voting took place to ask participants to vote each outcome as either “In” or “Out” using electronic keypads. Histograms and descriptive statistics were created live for each outcome during the meeting and displayed to the participants. Where the similar number of participants voted “In” or “Out,” issues were explored by discussion to determine the nature of the polarized response within the stakeholder groups. Dissenting views were actively sought and considered before voting was completed.

## Sample Size

There are no agreed methods to set the sample size for Delphi surveys or consensus meetings, and there is no requirement for a statistically representative sample [20]. Therefore, an opportunistic approach was used with the aim of obtaining approximately 100 respondents for both the professional and patient stakeholder groups for the survey and a smaller group in which discussion could take place in the consensus meetings.

## Data Analyses

After Round 1 of the survey, outcomes were categorized as “essential” and retained for Round 2 if they were rated between 7 and 9 by over 50% of respondents and between 1 and 3 by less than 15%. Outcomes not meeting these criteria for either patients or professionals were discarded. Mean scores were calculated for each retained outcome to form the feedback for Round 2. Round 2 responses were analyzed with stricter cut-off criteria, retaining outcomes rated between 7 and 9 by over 70% of respondents, and between 1 and 3 by less than 15%. There are no agreed methods for selecting cut-off criteria within Delphi studies and, therefore, the criteria were chosen after discussion within the writing group and collaborators within the COMET initiative.

The outcomes retained after Round 2 were considered in Phase 3 consensus meetings. During the meeting, each outcome was discussed, and voting took place that asked attendees to vote outcomes as “In,” “Out,” or during the patient meeting, “Unsure.” The “Unsure” category was included in the patient consensus meeting to ensure that participants understood the question. Voting was undertaken using electronic keypads to ensure anonymity, and no data were collected on participants who changed votes. The unsure items were rediscussed with further voting and discussion. All items retained from the patient and professional meetings were included in the final core set. There is no accepted definition of consensus in the literature. The overall approach in this study was to be inclusive so that outcomes of importance to participants were not inappropriately excluded from the COS. Therefore, outcome domains were only excluded if voted “In” by less than 33% of participants. There were deviations from this analysis. In the first professional consensus, meeting a more conservative approach was taken, because there was insufficient time for discussion. Domains were only excluded if voted “In” by less than 25% of participants. In addition, if consensus was not reached after two rounds of professional voting, a majority rules approach was taken.

## Results

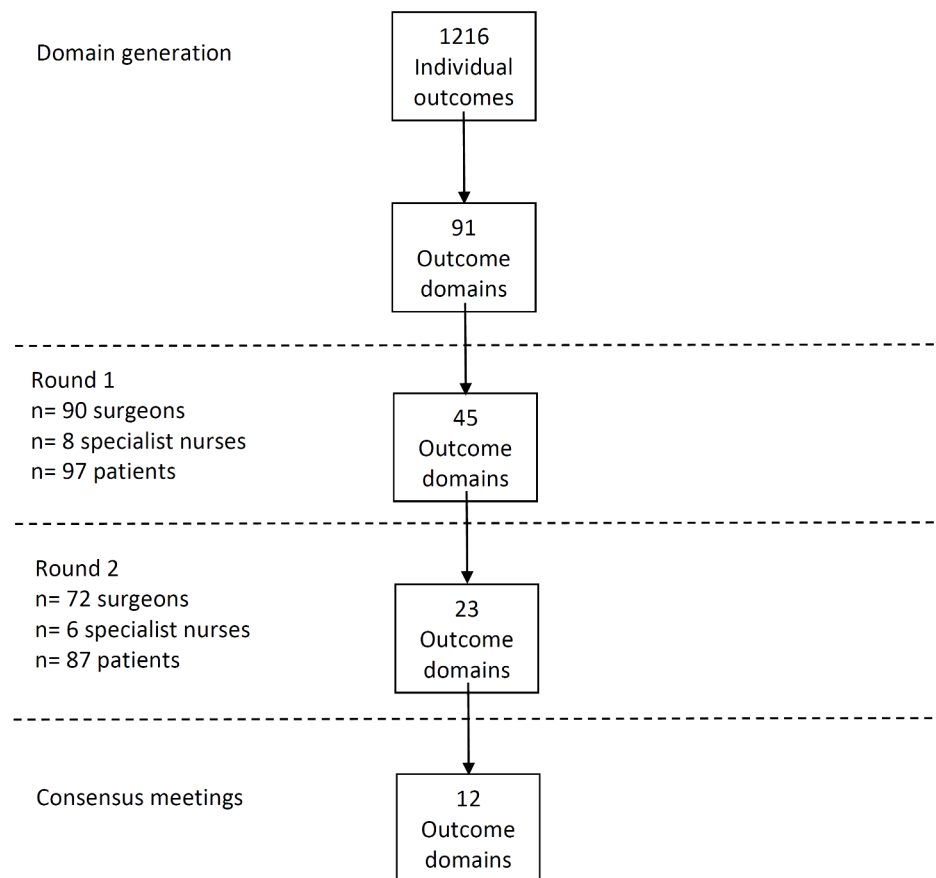
### Phase 1: Domain Generation

Review of all data sources identified 1,216 outcomes of CRC surgery that were grouped into 91 domains. The domains included outcomes about survival, recurrence, postoperative complications, and long-term quality of life. A summary of results is presented in [Fig 1](#).

### Phase 2: Delphi Process

A total of 81 CRC centers were sampled, of which 63 (78%) responded, including 90 surgeons and 8 clinical nurse specialists ([Table 1](#)). The centers represented all geographical regions of England and Wales, and caseloads averaged 117 major resections per year (range 38 to 275). Patient response rate was 97 out of 267 invited (36%). The patients’ age range was wide (29 to 87), sex ratio fairly equal (41 female, 42%), and similar numbers of patients had rectal (33, 35%), left (34, 35%), and right (30, 29%) colonic tumors. Many patients lived in areas of low deprivation, but there was an even distribution of basic and higher educational level. Health





**Fig 1. Flow diagram of Delphi process.**

doi:10.1371/journal.pmed.1002071.g001

professionals rated short-term technical outcomes of greatest importance in Round 1 including anastomotic leak, adequacy of resection margins, and perioperative mortality ([Table 2](#)). Although these issues were also rated as important to patients, patients gave a major priority to longer term outcomes such as survival, distant recurrence, and impact on longer term quality of life. A total of 45 domains met the criteria to be retained for Round 2 ([S1 Table](#)).

The response rate in Round 2 was 80% (78/98) for health professionals and 90% (87/97) for patients. The provision of feedback and more stringent cut-off criteria in Round 2 resulted in 23 domains being retained for consideration in the consensus meetings ([S2 Table](#)).

### Phase 3: Consensus Meetings

The two professional and one patient/caregiver consensus meetings were attended by 61, 35, and 14 participants, respectively. Professional demographic details were not completed as planned and are therefore missing. At the Tripartite colorectal conference, anonymized voting did not reach a consensus on domains for the core set in the allotted time. Eight domains were voted “Out” and were discarded. The remainder were considered polarized with support for inclusion of between approximately 40% and 60% ([Table 3](#)), and these were brought forward to the European Society of Coloproctology meeting. Initial voting in this second meeting identified four domains to be included into the core set, five to be discarded, and six to be discussed further. Follow-up voting reached a consensus on including an additional two domains

**Table 1. Participant characteristics.**

	Clinical centers	
	Responders (63)	Nonresponders (18)
Region (%)		
Northern England	14 (22)	5 (28)
Midland	8 (13)	0
Southeast England	22 (35)	10 (55)
Southwest England	9 (14)	0
Wales	10 (16)	3 (17)
Mean number of major colorectal resections (range) <sup>a</sup>	117 (38 to 275)	90 (29 to 210)
	Patients	
	Responders ( <i>n</i> = 97)	Nonresponders ( <i>n</i> = 170)
Mean age (range)	64 (29 to 87)	68 (29 to 88)
Female (%)	41 (42)	95 (56)
Cancer site (%)		
Rectum/anus	33 (35)	55 (32)
Left colon	34 (36)	46 (27)
Right colon	30 (29)	60 (36)
Unknown		9 (5)
IMD quintile (%) <sup>b</sup>		
1	5 (5)	27 (16)
2	13 (13)	38 (23)
3	20 (21)	24 (14)
4	20 (21)	41 (24)
5	39 (40)	23 (23)
Educational level (%)		
Basic	30 (32)	
Higher	34 (35)	
Undergraduate	16 (16)	
Postgraduate	6 (6)	
Not disclosed	11 (11)	
Marital status (%)		
Single/divorced	17 (18)	
Married/cohabiting	73 (75)	
Widowed	7 (7)	
Employment status (%)		
Employed	16 (17)	
Retired	58 (60)	
Seeking work	1 (1)	
Not working voluntarily	5 (5)	
Sickness leave	5 (5)	
Other	12 (12)	
Length of hospital stay (%)		
<2 weeks	80 (83)	
2–3 weeks	10 (10)	
3–4 weeks	3 (3)	
>4 weeks	4 (4)	

<sup>a</sup>Number of major cancer resections are defined by the UK National Bowel Cancer Audit 2012.

<sup>b</sup>IMD: Index of Multiple Deprivation as defined by the UK Office of National Statistics at lower layer Super Output Area level for the individual. Lower quintile equates to higher deprivation.

doi:10.1371/journal.pmed.1002071.t001



**Table 2. Top ten highest scored outcome domains after Round 1, by stakeholder group.**

Outcome domain	n (%) patients rating domain highly important <sup>a</sup>	Outcome domain	n (%) professionals rating domain highly important <sup>a</sup>
	<b>N = 97</b>		<b>N = 98</b>
Resection margins	88 (91)	Anastomotic leak	96 (99)
Stoma rate	84 (87)	Resection margins	93 (96)
Distant recurrence	81 (83)	Operative mortality	89 (92)
Recurrence	80 (82)	Conversion to open operation	88 (91)
Local recurrence	80 (82)	Distant recurrence	87 (90)
Nonprogression	80 (82)	Re-operation	87 (90)
Disease free interval	79 (81)	Local recurrence	86 (89)
Sphincter preservation	74 (76)	Recurrence	85 (88)
Lymph node yield	72 (74)	Lymph node yield	83 (86)
Survival	71 (73)	Length of hospital stay	83 (86)

<sup>a</sup>High importance is defined as scoring 7–9 on a nine-point Likert scale.

doi:10.1371/journal.pmed.1002071.t002

**Table 3. Voting on outcome domains to be included in the COS in the surgeon consensus meetings.**

Outcome domain	Tripartite voting N = 61 n (%) <sup>a</sup>		ESCP <sup>a</sup> Round 1 voting N = 35 n (%) <sup>*</sup>			ESCP <sup>a</sup> Round 2 voting N = 35 n (%) <sup>*</sup>		
	In	Consensus	In	Out	Consensus	In	Out	Consensus
Anastomotic leak	36 (59)	Vote again	23 (65)	7 (20)	Vote again	20 (57)	14 (40)	In
Surgical site infection	29 (48)	Vote again	12 (35)	12 (35)	Vote again	8 (23)	27 (77)	Out
Hemorrhage	13 (21)	Out	-	-	-	-	-	-
Visceral injury	9 (14)	Out	-	-	-	-	-	-
Conversion to open operation	15 (24)	Out	-	-	-	-	-	-
Venous thromboembolism	18 (29)	Vote again	4 (10)	28 (80)	Out	-	-	-
Bowel obstruction	16 (26)	Vote again	7 (21)	23 (67)	Out	-	-	-
Abandoning the operation	3 (5)	Out	-	-	-	-	-	-
Perioperative mortality	38 (62)	Vote again	33 (95)	2 (5)	In	-	-	-
Reoperation	43 (71)	Vote again	19 (55)	6 (17)	Vote again	13 (38)	20 (56)	Out
Stoma rate	25 (41)	Vote again	14 (40)	15 (42)	Vote again	16 (45)	17 (48)	Out
Stoma complications	16 (26)	Vote again	8 (23)	24 (68)	Out	-	-	-
Readmission to hospital	21 (34)	Vote again	15 (44)	18 (51)	Vote again	8 (23)	30 (77)	Out
Cancer recurrence	27 (45)	Vote again	32 (90)	2 (5)	In	-	-	-
Long-term survival	31 (50)	Vote again	28 (80)	5 (15)	In	-	-	-
Resection margins	20 (33)	Vote again	27 (78)	3 (9)	In	-	-	-
Lymph node yield	10 (16)	Out	-	-	-	-	-	-
Length of stay in hospital	13 (22)	Out	-	-	-	-	-	-
Quality of life	34 (56)	Vote again	19 (53)	9 (27)	Vote again	25 (71)	9 (26)	In
Sexual functioning	21 (34)	Vote again	7 (20)	21 (60)	Out	-	-	-
Physical functioning	12 (19)	Out	-	-	-	-	-	-
Fecal urgency	12 (19)	Out	-	-	-	-	-	-
Fecal incontinence	27 (44)	Vote again	9 (25)	18 (50)	Out	-	-	-

<sup>a</sup> European Society of Coloproctology

doi:10.1371/journal.pmed.1002071.t003

**Table 4. Voting on outcome domains to be included in the COS in the patient consensus meeting.**

Outcome domain	Round 1 voting n (%) <sup>a</sup>				Round 2 voting n (%) <sup>a</sup>			
	In	Out	Unsure	Consensus	In	Out	Unsure	Consensus
Anastomotic leak	9 (64)	2 (14)	3 (21)	Vote again	11 (79)	2 (14)	1 (7)	In
Surgical site infection	14 (100)	0	0	In	-	-	-	-
Hemorrhage	7 (50)	2 (14)	5 (36)	Vote again	1 (7)	10 (71)	3 (21)	Out
Visceral injury	8 (57)	3 (21)	3 (21)	Vote again	4 (29)	10 (71)	0	Out
Conversion to open operation	12 (86)	1 (7)	1 (7)	In	-	-	-	-
Venous thromboembolism	7 (50)	4 (29)	3 (21)	Vote again	4 (29)	8 (57)	2 (14)	Out
Bowel obstruction	5 (36)	5 (36)	4 (29)	Vote again	2 (14)	11 (79)	1 (7)	Out
Abandoning the operation	4 (29)	7 (50)	3 (21)	Out	-	-	-	-
Perioperative mortality	7 (50)	5 (36)	2 (14)	Vote again	4 (29)	10 (71)	0	Out
Reoperation	3 (21)	8 (57)	3 (21)	Out	-	-	-	-
Stoma rate	13 (93)	1 (7)	0	In	-	-	-	-
Stoma complications	6 (43)	6 (43)	2 (14)	Vote again	13 (93)	1 (7)	0	In
Readmission to hospital	5 (36)	8 (57)	1 (7)	Out	-	-	-	-
Cancer recurrence	9 (64)	4 (29)	1 (7)	In	-	-	-	-
Long-term survival	10 (71)	4 (29)	0	In	-	-	-	-
Resection margins	11 (79)	1 (7)	2 (14)	In	-	-	-	-
Lymph node yield	7 (50)	6 (43)	1 (7)	Vote again	0	13 (93)	1 (7)	Out
Length of stay in hospital	0	14 (100)	1 (7)	Out	-	-	-	-
Quality of life	12 (86)	1 (7)	1 (7)	In	-	-	-	-
Sexual functioning	7 (50)	3 (21)	4 (29)	Vote again	11 (79)	3 (21)	0	In
Physical functioning	10 (71)	4 (29)	0	In	-	-	-	-
Fecal urgency	10 (71)	1 (7)	3 (21)	In	-	-	-	-
Fecal incontinence	10 (71)	2 (14)	2 (14)	In	-	-	-	-

<sup>a</sup> Patients voted for whether each domain was in or out of the COS.

doi:10.1371/journal.pmed.1002071.t004

(Table 3). The composition of the final health professional core set of outcomes was ratified by a two-thirds majority.

In initial anonymized voting at the patient consensus meeting, ten domains were voted “In,” four “Out,” and nine considered for further debate (Table 4). Extensive discussion ensued, and it was recognized that some domains had overlapping content and meaning. “Physical functioning” was therefore grouped with “quality of life,” and “resection margins” was grouped with “survival.” Follow-up voting reached a consensus on including three more domains into the core set. Patient and professional COSs were then combined (Box 1). Discussions around perioperative mortality were interesting. Patients were aware that CRC surgery typically has a low operative mortality and did not feel it important to differentiate between early mortality and survival in the context of identifying the minimum set of core outcomes. It was excluded after two rounds of voting. Surgeons, conversely, felt perioperative mortality was an important marker of surgical (technical) success, and it was voted into the COS.

## Conclusions

This study has determined a COS to use in trials in CRC surgery. A wide range of sources including published studies and patient interviews were used to identify the initial long list that was reduced using consensus methods with professionals and patients to identify 23 domains

## Box 1. Final COS

### **Oncological outcomes:**

- Long-term survival
- Cancer recurrence
- Resection margins

### **Operative outcomes:**

- Anastomotic leak
- Perioperative survival
- Surgical site infection
- Stoma rates and complications
- Conversion to open operation (where appropriate)

### **Quality of life:**

- Physical function
- Sexual function
- Fecal incontinence
- Fecal urgency

of the greatest importance. Finally, consensus meetings with surgeons and patients and caregivers reconsidered the domains and voted on the final COS. It is recommended that all trials and other nonrandomized studies and audit undertaking of clinical evaluation of CRC surgery use this COS and further work to establish best instruments with which to measure these outcomes is underway.

It was not possible to identify other published COSs for CRC surgery. The COMET database has no other CRC COS development projects registered, although there is ongoing research to define a COS for anal cancer trials that may be conceptually similar [21]. A COS for use in all types of adult cancer treatment trials has been developed [22]. This generic cancer COS was developed with face-to-face consensus meetings with professionals who recommended that 12 symptoms be included in a COS (fatigue, insomnia, pain, anorexia, dyspnea, cognitive problems, anxiety, nausea, depression, sensory neuropathy, constipation, and diarrhea). It did not, however, survey patients' views, which are very important in the evaluation of treatments [23]. Conceptually, it has been argued that patients' views should be given at least equal if not greater importance over those of health professionals [24], and it is therefore unclear if this represents an appropriate COS. Furthermore, the scope of this COS encompasses all cancer treatments in adults. This broad remit may neglect details that are of specific importance to CRC patients or indeed patients undergoing surgery.

This study used robust consensus methodology and followed guidelines established by the COMET initiative to develop a COS, but there are some weaknesses. In Phase 1, the identification of large numbers of outcomes from primary data sources mandated the categorization

into domains. This introduces an element of subjectivity that was minimized through independent dual categorization, although there is the possibility that some outcomes may have been inappropriately grouped or separated. This is highlighted by the additional amalgamation of domains that occurred during the consensus meetings, where participants considered some domains unnecessarily detailed. In Phase 2, the scope of the Delphi process was limited to the UK before the COS development process was opened internationally to professionals in Phase 3. This was done to exclude the least important domains without the complexity of a multinational Delphi process; however, different domains may have been brought forward for discussion at the consensus meetings if the first round had included international participation. Participants at the professional consensus meeting did not report their country of residence as planned. This was not apparent until after the meeting had concluded, and it is therefore unclear as to the precise nationalities involved in the process. Further research is therefore needed to fully validate the COS more widely. This will include liaising with international organizations including the European Organisation for Research and Treatment of Cancer and the United States National Cancer Institute.

Another limitation is the numbers of participants involved in the process and response rates. In particular, the response rates from patients to the first questionnaire survey was low. The effect of this on the validity of the Delphi is unclear, because the purpose of the methodology is not to garner the views of a representative sample of stakeholders but to gain a consensus among a wide range of individuals with disparate opinions. In that respect, this study achieved wide diversity based on a priori patient characteristics. However, it is possible that patients not responding to the survey may have different opinions of the importance of each item to the responding group. Similarly, different professional groups, such as medical oncologists and radiologists, could have been recruited to bring a different perspective to the COS. The scope of the COS was, however, limited to surgery, and this guided the stakeholder involvement. It is important to expand the COS to include all treatment modalities in the future, at which point the involvement of other groups will be critical.

The scope of this COS was intentionally broad and included cancer of the colon and rectum. Many of the COS domains clearly traverse all colorectal surgery and include oncological outcomes such as survival, surgical outcomes such as anastomotic leak, and quality of life outcomes such as physical function. It is acknowledged, however, that patients have different experiences following surgery for colon or rectal cancer. Problems with sexual or bowel function, for example, are typically caused by the pelvic dissection and loss of reservoir associated with rectal surgery and are not usually associated with right-sided colonic surgery. Similarly, stoma formation is rare following right hemicolectomy. This issue was discussed at length in the professional consensus meetings and, although most participants agreed on a combined colorectal COS, some professionals still considered it unresolved. Nonetheless, feedback from patients suggested that these outcomes were important to measure in all colorectal studies, because the information was valued. In that respect, a patient undergoing right hemicolectomy may be concerned about the need for a stoma but reassured by a body of research demonstrating that stoma rates are low. Ultimately, the decision to have a combined colorectal COS was based on a patient-centered approach.

This study has defined which outcomes to measure in studies of CRC surgery. The next step is to identify how these outcomes should be measured in a valid, reliable, and acceptable way. This was not considered within this study because it is first necessary to assess the quality of potential outcome measures, a process that could not be undertaken until the COS domains were defined. One organization championing standards in measurement instruments is COS-MIN (CONsensus-based Standards for the selection of health Measurement Instruments) [25]. This group uses similar Delphi methods to agree on the taxonomy, terminology, and definition

of outcomes—a process that will be necessary to further the benefits of this COS. Another potential benefit of COSs is to provide evidence for use in clinical discussions with patients. Future research is required to examine how the COS can be included in clinical consultations to inform patient-centered decision making.

In conclusion, this study used health services research methodology to develop a COS for use in CRC surgical trials. It is now necessary to validate the use of this set in international research practice, with the aim of maximizing cross study comparisons, easing meta-analysis, and minimizing outcome reporting bias. Further work to identify recommended measures to use to assess each outcome is underway.

## Supporting Information

**S1 Table. Patient and professional scoring of outcome domains in Round 1, including details of which domains were retained by patients, professionals, and overall.** Domains were retained if rated of high importance by over 50% of respondents and low importance by less than 15% of respondents. Domains were retained overall if they were retained by either stakeholder.

(DOCX)

**S2 Table. Patient and professional scoring of outcome domains in Round 2, including details of which domains were retained by patients, professionals, and overall.** Domains were retained if rated of high importance by over 70% of respondents and low importance by less than 15% of respondents. Domains were retained overall if they were retained by either stakeholder.

(DOCX)

## Acknowledgments

We would like to thank Barry Main, Neil Smart, and Katherine Gash for their help running the consensus meetings, Claudette Blake for her administrative support throughout the whole project, and James Jones and George Smith for their input as patient representatives.

## Author Contributions

Wrote the first draft of the manuscript: AGKM RNW RM JR JMB. Contributed to the writing of the manuscript: AGKM RNW ROF RM JR AMP MGT AR DM SJD JB KNLA STB JMB. Agree with the manuscript's results and conclusions: AGKM RNW ROF RM JR AMP MGT PAS AR AO DM RK DGJ RHu RHs SJD MGC MC JB KNLA STB JMB. All authors have read, and confirm that they meet, ICMJE criteria for authorship.

## References

1. Hirsch BR, Califf RM, Cheng SK, Tasneem A, Horton J, Chiswell K, et al. Characteristics of oncology clinical trials: insights from a systematic analysis of ClinicalTrials.gov. *JAMA internal medicine*. 2013; 173(11):972–9. doi: [10.1001/jamainternmed.2013.627](https://doi.org/10.1001/jamainternmed.2013.627) PMID: [23699837](https://pubmed.ncbi.nlm.nih.gov/23699837/).
2. Meher S, Alfrevic Z. Choice of primary outcomes in randomised trials and systematic reviews evaluating interventions for preterm birth prevention: a systematic review. *BJOG: an international journal of obstetrics and gynaecology*. 2014; 121(10):1188–94; discussion 95–6. doi: [10.1111/1471-0528.12593](https://doi.org/10.1111/1471-0528.12593) PMID: [24571433](https://pubmed.ncbi.nlm.nih.gov/24571433/).
3. Tsiachlakis A, O'Brien K. Do orthodontic research outcomes reflect patient values? A systematic review of randomized controlled trials involving children. *American journal of orthodontics and dentofacial orthopedics: official publication of the American Association of Orthodontists, its constituent societies,*

- and the American Board of Orthodontics. 2014; 146(3):279–85. doi: [10.1016/j.ajodo.2014.05.022](https://doi.org/10.1016/j.ajodo.2014.05.022) PMID: [25172249](https://pubmed.ncbi.nlm.nih.gov/25172249/).
4. Rodgers S, Brealey S, Jefferson L, McDaid C, Maund E, Hanchard N, et al. Exploring the outcomes in studies of primary frozen shoulder: is there a need for a core outcome set? *Qual Life Res*. 2014; 23(9):2495–504. doi: [10.1007/s11136-014-0708-6](https://doi.org/10.1007/s11136-014-0708-6) PMID: [24817317](https://pubmed.ncbi.nlm.nih.gov/24817317/).
5. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ*. 2010; 340:c365. doi: [10.1136/bmj.c365](https://doi.org/10.1136/bmj.c365) PMID: [20156912](https://pubmed.ncbi.nlm.nih.gov/20156912/).
6. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials*. 2012; 13:132. doi: [10.1186/1745-6215-13-132](https://doi.org/10.1186/1745-6215-13-132) PMID: [22867278](https://pubmed.ncbi.nlm.nih.gov/22867278/); PubMed Central PMCID: PMC3472231.
7. US Department of Health Human Services FDA. 1999.
8. Guideline on Clinical Investigation of Medicinal Products other than NSAIDs for Treatment of Rheumatoid Arthritis. The European Agency for the Evaluation of Medicinal Products [Internet]. 2003. [http://www.emea.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2009/09/WC500003439.pdf](http://www.emea.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003439.pdf).
9. Khan K. The CROWN Initiative: journal editors invite researchers to develop core outcomes in women's health. *Midwifery*. 2014; 30(12):1147–8. doi: [10.1016/j.midw.2014.10.001](https://doi.org/10.1016/j.midw.2014.10.001) PMID: [25434973](https://pubmed.ncbi.nlm.nih.gov/25434973/).
10. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for Outcome Measures in Rheumatology. *The Journal of rheumatology*. 1998; 25(2):198–9. PMID: [9489805](https://pubmed.ncbi.nlm.nih.gov/9489805/).
11. McGrath PJ, Walco GA, Turk DC, Dworkin RH, Brown MT, Davidson K, et al. Core outcome domains and measures for pediatric acute and chronic/recurrent pain clinical trials: PedIMMPACT recommendations. *The journal of pain: official journal of the American Pain Society*. 2008; 9(9):771–83. doi: [10.1016/j.jpain.2008.04.007](https://doi.org/10.1016/j.jpain.2008.04.007) PMID: [18562251](https://pubmed.ncbi.nlm.nih.gov/18562251/).
12. Devane D, Begley CM, Clarke M, Horey D, OB C. Evaluating maternity care: a core set of outcome measures. *Birth*. 2007; 34(2):164–72. doi: [10.1111/j.1523-536X.2006.00145.x](https://doi.org/10.1111/j.1523-536X.2006.00145.x) PMID: [17542821](https://pubmed.ncbi.nlm.nih.gov/17542821/).
13. World Health O. Colorectal cancer Estimated Incidence, Mortality and Prevalence Worldwide in 2012 2012 [cited 2012 23/12/2015]. [http://globocan.iarc.fr/Pages/fact\\_sheets\\_cancer.aspx?cancer=colorectal](http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx?cancer=colorectal).
14. Sullivan R, Alatise OI, Anderson BO, Audisio R, Autier P, Aggarwal A, et al. Global cancer surgery: delivering safe, affordable, and timely cancer surgery. *Lancet Oncol*. 2015; 16(11):1193–224. doi: [10.1016/S1470-2045\(15\)00223-5](https://doi.org/10.1016/S1470-2045(15)00223-5) PMID: [26427363](https://pubmed.ncbi.nlm.nih.gov/26427363/).
15. Kuhry E, Schwenk W, Gaupset R, Romild U, Bonjer HJ. Long-term results of laparoscopic colorectal cancer resection. *Cochrane Database of Systematic Reviews* [Internet]. 2008; (2: ). <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003432.pub2/abstract> <http://onlinelibrary.wiley.com/store/10.1002/14651858.CD003432.pub2/asset/CD003432.pdf?v=1&t=i87721gi&s=25558cf9c9ff7916b003eb2494dbc515efaf81abf>.
16. Schwenk W, Haase O, Neudecker Jens J, Müller Joachim M. Short term benefits for laparoscopic colorectal resection. *Cochrane Database of Systematic Reviews* [Internet]. 2005; (2: ). <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003145.pub2/abstract> <http://onlinelibrary.wiley.com/store/10.1002/14651858.CD003145.pub2/asset/CD003145.pdf?v=1&t=i8771wha&s=5dd1f7c09f5a835a98ee37ce5c1227d75927ce18>.
17. Whistance RN, Forsythe RO, McNair AG, Brookes ST, Avery KN, Pullyblank AM, et al. A systematic review of outcome reporting in colorectal cancer surgery. *Colorectal Dis*. 2013; 15(10):e548–60. doi: [10.1111/codi.12378](https://doi.org/10.1111/codi.12378) PMID: [23926896](https://pubmed.ncbi.nlm.nih.gov/23926896/).
18. McNair A, Whistance RN, Forsythe RO, Rees J, Jones JE, Pullyblank AM, et al. Synthesis and summary of patient-reported outcome measures (PROMs) to inform the development of a core outcome set in colorectal cancer surgery. *Colorectal Dis*. 2015. doi: [10.1111/codi.13021](https://doi.org/10.1111/codi.13021) PMID: [26058878](https://pubmed.ncbi.nlm.nih.gov/26058878/).
19. Macefield RC, Jacobs M, Korfage IJ, Nicklin J, Whistance RN, Brookes ST, et al. Developing core outcomes sets: methods for identifying and including patient-reported outcomes (PROs). *Trials*. 2014; 15:49. doi: [10.1186/1745-6215-15-49](https://doi.org/10.1186/1745-6215-15-49) PMID: [24495582](https://pubmed.ncbi.nlm.nih.gov/24495582/); PubMed Central PMCID: PMC3916696.
20. Powell C. The Delphi technique: myths and realities. *J Adv Nurs*. 2003; 41(4):376–82. PMID: [12581103](https://pubmed.ncbi.nlm.nih.gov/12581103/).
21. COMET Initiative. COMET database 2015 [20/05/2015]. <http://www.comet-initiative.org/studies/search>.
22. Reeve BB, Mitchell SA, Dueck AC, Basch E, Cella D, Reilly CM, et al. Recommended patient-reported core set of symptoms to measure in adult cancer treatment trials. *J Natl Cancer Inst*. 2014; 106(7). doi: [10.1093/jnci/dju129](https://doi.org/10.1093/jnci/dju129) PMID: [25006191](https://pubmed.ncbi.nlm.nih.gov/25006191/); PubMed Central PMCID: PMC4110472.

23. Main BG, Blencowe N, Williamson PR, Blazeby JM. RE: Recommended Patient-Reported Core Set of Symptoms to Measure in Adult Cancer Treatment Trials. *Journal of the National Cancer Institute*. 2015; 107(4). doi: [10.1093/jnci/dju506](https://doi.org/10.1093/jnci/dju506)
24. Main BG, Strong S, McNair AG, Falk SJ, Crosby T, Blazeby JM. Reporting outcomes of definitive radiation-based treatment for esophageal cancer: a review of the literature. *Dis Esophagus*. 2014. doi: [10.1111/dote.12168](https://doi.org/10.1111/dote.12168) PMID: [24438540](https://pubmed.ncbi.nlm.nih.gov/24438540/).
25. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010; 63(7):737–45. doi: [10.1016/j.jclinepi.2010.02.006](https://doi.org/10.1016/j.jclinepi.2010.02.006) PMID: [20494804](https://pubmed.ncbi.nlm.nih.gov/20494804/).